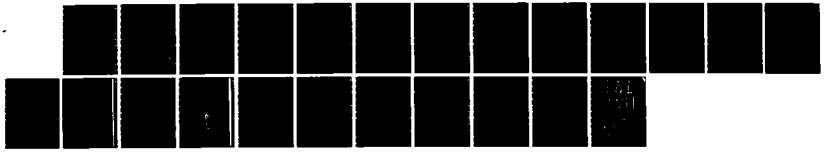
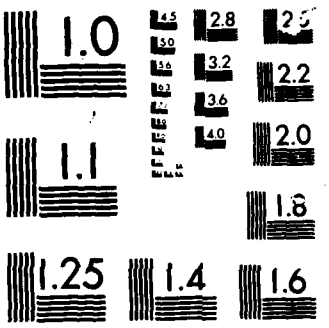
AD-A166 787 COMPUTER-AIDED FABRICATION SYSTEM IMPLEMENTATION(U) 1/1
MASSACHUSETTS INST OF TECH CAMBRIDGE P PENFIELD ET AL.
31 MAR 86 N00014-85-K-0213

UNCLASSIFIED F/G 9/2 NL





MICROCOPY CHART

(12)

AD-A166 787

COMPUTER-AIDED FABRICATION SYSTEM IMPLEMENTATION

Semiannual Technical Report for the period October 1, 1985 to March 31, 1986

Massachusetts Institute of Technology
Cambridge, MA 02139

Principal Investigators:
Paul Penfield, Jr. (617) 253-2506
Stanley B. Gershwin (617) 253-2149
Dimitri A. Antoniadis (617) 253-4693
Donald E. Troxel (617) 253-2570

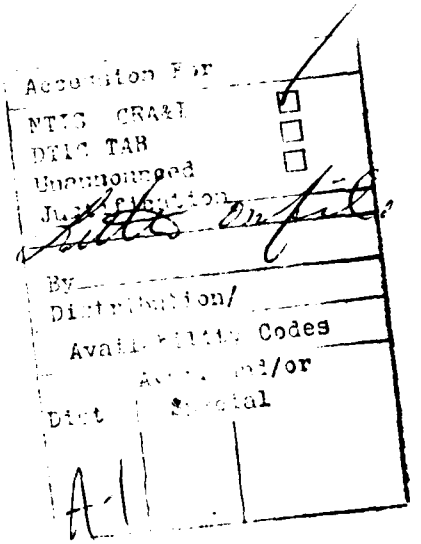
This research was sponsored by Defense Advanced Research Projects Agency (DoD), through the Office of Naval Research under Contract N00014-85-K-0213.

TABLE OF CONTENTS

Selected Publications (starting after page 8)

T.-L. Tung and D. A. Antoniadis, "A Boundary Integrated Equation Approach to Oxidation Modeling," IEEE Transactions on Electron Devices, vol. ED-32, no. 10, pp. 1954-1958, October 1985.

C. Lozinski and S. B. Gershwin, "Dynamic Production Scheduling in Computer-Aided Fabrication of Integrated Circuits," to be presented at the IEEE Robotics and Automation Conference, San Francisco, CA, April 7-10, 1986.



RESEARCH OVERVIEW

This report covers the period from October 1, 1985 through March 31, 1986. The research discussed here is described in more detail in several published and unpublished reports cited below.

During this period, the installation of the central computing machine for the MIT CAF work, known as CAF (or more formally, as caf.mit.edu) and all the necessary software was substantially completed. The machine is a VAX 11/785. It has a large cadre of users.

A rudimentary CAF system, known as CAFE (Computer-Aided Fabrication Environment), has been implemented. So far this system provides an electronic personal notebook, machine instructions, and equipment status reports. Four fabrication facilities are being served, at least in part. Even though crude, CAFE has been in constant use.

The requirements of a scheduling program are being studied. One surprising result so far is the importance of setup operations in scheduling some equipment.

Two new process-simulation modules have been developed. One is for a focused ion beam operation used for implanting, and the other is for nonuniform oxidation. Improved models were inserted into the simulators SUPREM and MINIMOS.

Keywords: computer aided manufacturing; industrial production; computer aided design

SCHEDULING

Research during this period focused on three activities: studying the integrated-circuit fabrication process at a systems level, formulating a mathematical model of an integrated-circuit fabrication facility, and developing an electronic sign-up sheet for scheduling the laboratory.

The first study is to define the scheduling problem. This involves studying VLSI technology, the construction of the MIT Integrated Circuits Laboratory, and user requirements. It was discovered that a few key elements of the problem may have dramatic impact on the final schedule. A setup is required when there is an important change in type of operation at a machine, and can be due to changes of gases, water temperatures, or other required materials. Setups complicate scheduling because machines that are undergoing such changeovers are not available for productive operations. There is an incentive to put off changeovers as long as possible because when they take place often, production capacity is reduced. On the other hand, if changeovers are performed too infrequently, some lots will be delayed a long time.

Randomness is due to machine failures and unpredictable demand on the facility. In a university setting, the mix of chips going through the facility has very little consistency from week to week. We will have to group the demands for scheduling purposes, and use guesses based on historical data. The presence of randomness means that any schedule must have some idle time in it, so that a perturbation does not propagate indefinitely. It also means that there must be some means for rescheduling after events like machine failures.

We are also investigating the differences between a university laboratory and a commercial production system. In a commercial facility, machine failures are no less important, but there is a more predictable demand on the facility. The CAF/Scheduling group visited DEC in Hudson, Massachusetts on February 11, 1986 to observe their production system.

The second activity is to develop a scheduling model of the system. Developing such a model requires identifying scheduling objectives, decomposing the problem, solving each of the subproblems, developing a database to store the required information, and testing the solution on the facility. Scheduling is complicated by setups and randomness, and various sources of both are being studied. The main focus of the analytic activity is the setup issue. The approach now being explored views setups as events that take place at a lower frequency than machine operations and most other events. A slow-time-scale model is being explored that accounts for the time that the system spends in each setup configuration, as well as the time spent in changing configurations. The impact of setups on capacity is modeled explicitly.

Remaining work on this formulation includes modeling the scheduling objective, and then translating the setup frequencies that are calculated this way into actual times at which to perform the changes in configurations.

The purpose of this activity is to develop a simple Version 1 scheduler which will not use the sophisticated ideas that are described above, but which can be implemented and used relatively quickly. The existing paper system is being designated as Version 0, and the first electronic system as Version 1. Great consideration is being given to ease of use. It will be designed to emulate the paper sign-up sheet in the old facility, with no additional features except for the fact that it can be accessed from any terminal.

MODULAR PROCESS

The focus of this project is the development of methods for IC fabrication process and device design in the context of a Computer Aided Fabrication (CAF) research laboratory. The main thrust of our work centers on the development of what we call the MASTIF workstation (MIT Analysis and Simulation Tools for IC Fabrication). This is a menu- and window-oriented program written in C and FORTRAN which provides methodology and uniform software structure for the connection of process and device-design CAD tools. MASTIF currently includes a facility for incremental development and version management of a process description, management mechanisms for definition of physical cross sections deriving from the overall process description, and an interactive graphics interface and data interchange for process and device simulators (SUPREM-III and MINIMOS). With MASTIF the user can effectively develop and evaluate a fabrication process via a single integrated workstation.

In parallel with the development of MASTIF, work is under way in the area of process and device modeling. Extensive use of SUPREM-III has revealed several shortcomings of the program. Perhaps the most significant improvement that we have implemented in an experimental fashion is the inclusion of dynamic clustering for arsenic atoms in silicon. Similarly, we have scrutinized, found inaccurate, and improved significantly the electron-mobility model in the device simulator MINIMOS.

Two new process-simulation modules have been developed during this period. The first is used for the generation of as-implanted and diffused-doping profiles generated by means of Focused Ion Beam (FIB) implantation. This is in support of a separate FIB applications research program also supported by DARPA.

The second process-simulation module is still in an experimental form. It allows the simulation of nonuniform silicon oxidation. The calculated results are the shape of the grown SiO_2 and the stress exerted during the process at the Si/SiO_2 interface. A boundary-integral-equation method has been developed for the numerical solution of the problem. Because of the many uncertainties in physical parameters for grown SiO_2, the present simulator allows the investigation of three different oxide-movement behaviors: elastic, viscous, and visco-elastic. Work is under way to establish, by comparison with experimental data, the range of validity of these different models.

CAF SYSTEM STRUCTURE

Prior to the beginning of this reporting period, an initial, rudimentary CAF system was designed and implemented. This included a personal notebook and data structures for the upcoming fabrication equipment installation. This was implemented on a borrowed VAX 11/750 and some of the engineering staff started to use the lab notebook and editor. RS232 cables to the offices and labs were installed and some were checked and connected to terminal concentrator ports. Three concentrators were assembled, connected to an MIT Ethernet, and were operational. The VAX 11/785 hardware was installed.

Since then, we have completed the computer hardware installation, evaluated and purchased terminals for both users and for use in the clean rooms, installed Unix version 4.3BSD, integrated local networking software, adapted mail software to our local needs, ported several applications programs to our CAF computer, instituted a weekly back-up procedure, substantially improved our terminal concentrators, created user documentation, interfaced to the MIT physical plant computer which monitors a large number of sensing points in our facility, initiated some student projects, initiated plans for including the teaching laboratories in our developing CAF system, acquired both documentation and software for the present Berkeley CAF system, initiated acquisition of additional main memory and disk storage, and expanded our user base to include virtually all of the faculty, students, and technical staff associated with our fabrication facility.

Our present computer hardware consists of a VAX 11/785 with 8 Mbytes of main memory, two 450 Mbyte disks, GCR tape, 1600-cpi tape, laser printer, four phone lines and modems, and an Ethernet port to the MIT network. Additionally there are seven 11/73-based terminal concentrators, each of which has 32 RS232 ports. RS232 cables have been installed throughout our building, and numerous user terminals and PCs have been connected.

We have completed an extensive evaluation of terminals for use in the clean rooms, and by the technical staff and other CAF users. The information gleaned during our selection process has been communicated to other building occupants for their consideration. We chose Ergo 301 terminals as they had the best compromise considering cost, video quality, compatibility with existing software, and the fact that they do not have fans, which allows their use in the clean room areas.

We installed the latest version of UNIX, Berkeley 4.3BSD, and have integrated local networking software and have adapted the mail software to enable local as well as remote communications. Our official host name is caf.mit.edu; locally the computer is known simply as caf.

Several large applications programs have been successfully ported to run on mit-caf, including MAGIC and PISCES. The success in porting these programs has resulted in greater use of the caf system, so much that we will soon need increased primary memory and user disk space.

A back-up regimen has been initiated with full dumps taken every month and incremental dumps every week. This is intended primarily as protection against catastrophic disk failure.

Experiments with our terminal concentrators revealed service deficiencies when all ports were active. The cause was traced to inadequate use of the available buffer memory by the SWITCH software. This has now been corrected and satisfactory performance is achieved with 32 active lines per concentrator. We are reasonably confident that we could implement 48 lines per concentrator but are not sure that satisfactory response times would be preserved if all 48 lines were active at once. We have more physical memory than can now be used, probably enough to service twice the number of lines per concentrator. However, it is doubtful that the rather substantial software rewrite required to use this memory is worth the effort.

We have automated the procedures for applying for and approving new computer accounts. An program named "open" has been written which enables prospective users to make requests for accounts and provide the required information. Mail is sent to an "approver" after an account has been requested and new accounts, including appropriate initialization files and directories, are automatically created after approval.

We have also been active in the generation of documentation for CAF users. User documentation now covers the use of terminal concentrators, obtaining accounts, obtaining documentation, using UNIX and EMACS, using Kermit with IBM PCs, and using the nroff text formatter with the -me macro package. Manual entries for the Cshell with command completion and Mail utility have also been made available for distribution. We have initiated a project to define laser printer hardware and software to provide both typesetting and hard copy of arbitrary graphics.

We have interfaced to the MIT physical plant computer which monitors a large number of sensing points such as resistance of DI water, etc., in our facility. We capture all alarms printed and have provided a mechanism for remote query of the status of the monitored points. Our facilities personnel can now interrogate the monitoring computer from either office or home and efficiently direct corrective action.

We have initiated a plan to expand the availability of the CAF system to the integrated-circuit teaching laboratory. This will provide us with another active test bed for the CAF system, and will also allow indoctrination of entering graduate students (our present practice is to require new students to complete an intensive, short version of the undergraduate teaching laboratory subject).

We have advertised to our students several projects concerned with integrating measuring instruments into the CAF system. One student has started work on a project relating to a CV measuring instrument. Another project has been defined concerning software for a scanning thermometer intended for use in monitoring photolithography process parameters.

We have acquired both documentation and software for the present Berkeley CAF system. We are in the process of installing a test version of this software and will study the integration of this software into our system or the integration of our system into the Berkeley software.

We now have a substantial number of CAF users. The system response time remains quite good, but we have had to make plans to augment primary memory and available disk space. Our primary future activity is to further the development of the CAF applications software and its use in the installation and operation of equipment to be installed.

PUBLICATIONS LIST

T.-L. Tung and D. A. Antoniadis, "A Boundary Integrated Equation Approach to Oxidation Modeling," IEEE Transactions on Electron Devices, vol. ED-32, no. 10, pp. 1954-1958, October 1985.

D. Boning and D. A. Antoniadis, "MASTIF - A Workstation Approach to Fabrication Process Design," Digest of Technical Papers, IEEE International Conference on CAD, ICCAD-85, Santa Clara, CA, pp. 280-282, November 18-21, 1985.

C. Lozinski and S. B. Gershwin, "Dynamic Production Scheduling in Computer-Aided Fabrication of Integrated Circuits," to be presented at the IEEE Robotics and Automation Conference, San Francisco, CA, April 7-10, 1986.

TALKS WITHOUT PROCEEDINGS

P. Penfield, Jr., "Computer-Aided Fabrication and Other VLSI Research at MIT," IBM Corporation, Yorktown Heights, NY, March 21, 1986.

D. Boning and D. A. Antoniadis, "MASTIF - A Workstation Approach to Fabrication Process Design," MIT Fall 1985 VLSI Research Review, Cambridge, MA, December 16, 1985.

A Boundary Integral Equation Approach to Oxidation Modeling

THYE-LAI TUNG, STUDENT MEMBER, IEEE, AND DIMITRI A. ANTONIADIS, SENIOR MEMBER, IEEE

Abstract—**Thermal oxidation of silicon involves the diffusion of oxidant molecules from the gas-oxide interface to the oxide-silicon interface, and the transport of newly formed oxide away from the latter. Under suitable formulations these two processes can be shown to be boundary-value problems of harmonic and biharmonic nature. Based on these properties, a boundary integral equation method (BIEM) has been developed for modeling oxidation. This method achieves simplicity and efficiency by solving a two-dimensional problem using line integrals on the boundaries. The use of source distributions as intermediary solutions facilitates direct calculations of a wide variety of boundary parameters.**

Introduction

In local oxidation of silicon (LOCOS) process, the shape of the oxide undergoes drastic transformation because of nonuniform oxide growth rate. A numerical method that can model such change with ease is desirable for comprehensive IC fabrication simulation. Methods based on integral equations have previously been explored for modeling oxidation [1], [2]. Compared to conventional methods such as finite difference and finite element, they have achieved some advantages, for instance, larger tolerance on the deformations of the mesh; nevertheless they still require calculations within the oxide bulk.

This paper reports a Boundary Integral Equation Method (BIEM) for modeling thermal oxidation in two dimensions. Our method relies on Green's theorem for harmonic and biharmonic problems, and, unlike all previous approaches, obtains solutions solely from line integrals along the oxide boundaries, achieving reduction in dimensionality. Because it does not require a mesh to subdivide the oxide bulk, it is simple and efficient. We only have to deal with the boundaries, precisely what is needed in many applications.

Physical Models

The two constituent processes of thermal oxidation are the diffusion of oxidant and the motion of oxide. In the oxidant transport process, oxidant molecules diffuse from the ambient, through the oxide bulk, to the oxide-silicon interface to react with silicon. Due to its volume expansion, newly created oxide partly fills the void left by consumed silicon, and partly displaces and pushes existing oxide towards the surface. At sufficiently high temperatures, oxide behaves like a viscous fluid, thus its motion can be studied as a fluid flow problem.

Models used in our method are identical to those in Chin's [1]. Following the approach of the Deal-Grove one-dimensional oxidation model [3], the oxidant diffusion is treated as a steady-state process with uniform diffusivity. The oxidant flux $\vec{F}$ is given by

$$\vec{F} = -D\nabla C \tag{1}$$

where D and C are the diffusion coefficient and concentration of oxidant. The steady-state assumption requires

$$\nabla \cdot \vec{F} = 0 \tag{2}$$

this immediately suggests that C satisfies the Laplace's equation if D is a constant:

$$\nabla^2 C = 0. \tag{3}$$

The oxide is modeled as a viscous incompressible fluid described by the creeping flow and continuity equations:

$$\mu\nabla^2\vec{u} = \nabla P \tag{4}$$

$$\nabla \cdot \vec{u} = 0 \tag{5}$$

where μ, $\vec{u}$, and P are the viscosity, velocity and pressure of oxide respectively. The viscous fluid approximation is valid as long as the oxidation temperature is above 960°C. Below the critical temperature, oxide exhibits strong elastic characteristics [4].

Incompressible flow in two dimensions can be described by a stream function Ψ, which is related to $\vec{u}$ by

$$\vec{u} = \frac{\partial\Psi}{\partial y}\hat{x} - \frac{\partial\Psi}{\partial x}\hat{y} \tag{6}$$

where $\hat{x}$ and $\hat{y}$ are unit vectors in the x and y directions. For creeping flow, Ψ satisfies the biharmonic equation:

$$\nabla^4\Psi = 0 \tag{7}$$

where ∇^4 is the Bi-Laplace operator:

$$\nabla^4 = \frac{\partial^4}{\partial x^4} + 2\frac{\partial^4}{\partial x^2\partial y^2} + \frac{\partial^4}{\partial y^4}. \tag{8}$$

Normally a vector, the vorticity ω is treated here as a scalar since it is always normal to the plane in two-dimensional problems; it is given by

$$\omega = \frac{\partial u_y}{\partial x} - \frac{\partial u_x}{\partial y} \tag{9}$$

Manuscript received January 29, 1985; revised June 20, 1985. This work was supported by the General Electric Company and the Defense Advanced Research Projects Agency.

The authors are with Massachusetts Institute of Technology, Cambridge, MA 02139.

0018-9383/85/1000-1954$01.00 © 1985 IEEE

where the subscripts x and y denote the x and y components of vector $\vec{u}$. It follows from (6) that ω is related to Ψ by

$$\omega = -\nabla^2\Psi. \tag{10}$$

An important feature in this biharmonic formulation is that P and ω are related by the Cauchy–Rieman equations [5]:

$$\frac{\partial\omega}{\partial x} = \frac{\partial(P/\mu)}{\partial y}$$

$$\frac{\partial\omega}{\partial y} = -\frac{\partial(P/\mu)}{\partial x}. \tag{11}$$

This conjugate harmonic relationship enables us to obtain P from Ψ with the following operator:

$$\overline{\nabla}^2 \tag{12}$$

dubbed as "del-square conjugate" operator. This operator takes the Laplacian of Ψ to produce $-\omega$, and conjugates the result to yield $-P/\mu$.

With the principal components, namely, $\vec{u}$ and P, readily available from Ψ, all types of boundary conditions can be modeled. A list of those boundary conditions is given in the Appendix.

As a masking material impermeable to most oxidants, silicon nitride makes LOCOS possible. In addition to blocking the supply of oxidant species from the ambient, it acts as barrier to the flow of oxide. A simplified beam bending equation described in [9] had been adopted to model bending of the silicon nitride mask, but no satisfactory results were obtained. Thus for the time being the silicon nitride mask is assumed to be sufficiently thin, that is less than 300 Å, and flexible that it does not exert any normal force on the oxide. It should be pointed out that different silicon nitride models can be incorporated into our program readily because they only enter as boundary conditions.

Solution Method

Harmonic and biharmonic problems can be solved efficiently using the integral form of the governing equations. Both the oxidation diffusion and oxide flow problems are modeled using the source distribution technique in which the solutions are expressed in terms of source distributions on the boundaries [6], [7]:

$$C(\vec{p}) = \int_{\partial B} \sigma(\vec{q})K(\vec{p} - \vec{q})\,dq \tag{13}$$

$$\Psi(\vec{p}) = \int_{\partial B} \sigma_1(\vec{q})K_1(\vec{p} - \vec{q}) + \sigma_2(\vec{q})K_2(\vec{p} - \vec{q})\,dq \tag{14}$$

where σ's and K's are the sources and the kernels, respectively, and ∂B denotes the boundaries. A simple analogy for the harmonic problem is the potential distribution in an enclosed space. In this case, σ is just positive and negative charges, and K the potential distribution due to a point charge. Charges are arranged in such a way that their net potential contribution matches the prescribed boundary condition, and this gives the solution for the whole region. A detailed report of the biharmonic method, applicable to general creeping flow problems, will appear in another article.

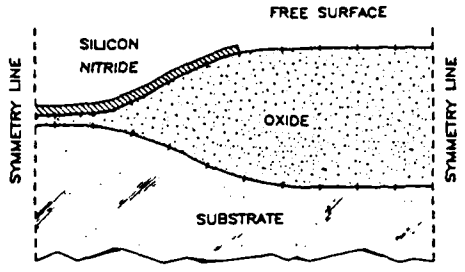


Fig. 1. Boundary segmentation for a SemiRox structure. Segments are indicated by tic marks. Three different types of boundary are present: free surface where the oxide is in contact with the gas, silicon nitride mask, and oxide–silicon interface.

As with simulations of other physical continua, we limit our oxidation modeling to a small region of a wafer due to practical considerations. Reflective boundaries must be introduced, one on the left, one on the right, to enclose the region. However, artificial walls may not be arbitrarily used in integral equation technique, for the sharp corners created between the natural and artificial boundaries will introduce large errors in the numerical solutions. To circumvent this problem, simulation structures are mirror imaged to become periodic symmetric. In this case, the symmetry lines represent reflecting boundaries. Dimensions are scaled to be 2π periodic and K is modified to be periodic. As an example, (13) is rewritten as

$$\begin{aligned} C(\vec{p}) &= \int_{\partial B_l} \sigma(\vec{q}) \sum_{n=-\infty}^{\infty} K(2\pi n\hat{\imath};\ \vec{p} - \vec{q})\,dq \\ &\quad + \int_{\partial B_u} \sigma(\vec{q}) \sum_{n=-\infty}^{\infty} K(2\pi n\hat{\imath} + \vec{p} - \vec{q})\,dq \\ &= \int_{\partial B_l} \sigma(\vec{q})\tilde{K}(\vec{p} - \vec{q})\,dq \\ &\quad + \int_{\partial B_u} \sigma(\vec{q})\tilde{K}(\vec{p} - \vec{q})\,dq \end{aligned} \tag{15}$$

where ∂B_l is a period of the lower bundary, and ∂B_u of the upper boundary. The tilde symbol denotes periodicity, that is $\tilde{K}(\vec{p}) = \tilde{K}(\vec{p} + 2\pi n\hat{\imath})$ for any integer n. Forms of the periodic kernels are given in the Appendix.

The boundaries are divided into straight line segments, as illustrated in Fig. 1. Typically there are 20 to 35 segments per half period for each of the top and the bottom boundaries. The lower limit is due to accuracy consideration, and the upper limit is constrained by computation speed. After the boundaries are segmented, the source distributions on those segments are approximated with step functions, that is the source strength is constant along the length of a segment. The midpoints of the segments are chosen to be the collocation points (test points) where boundary conditions are imposed. We recall that bound-

ary parameters are obtained by applying differentiation operators on C and Ψ. According to (13) and (14), these operators in turns operate on K, K_1, and K_2. Thus for the oxidant diffusion problem, the boundary condition at boundary segment i is given by

$$b_i = \sum_{j=1}^{N} \sigma_j \int_{\partial B_j} \mathcal{L}_i[K(\vec{p}_i - \vec{q}_j)]\, dq_j \qquad (16)$$

where b_i, $\mathcal{L}_i$, $\vec{p}_i$ are the boundary condition, associated operator, and midpoint of segment i, respectively, and N is the number of segments.

Each integral within the summation of (16) is obtained from either a fourth-order or a sixth-order Newton-Cotes quadrature [8]. The quadrature scheme however fails when the source and collocation points coincide and produce a singular integrand. In this case the kernels are converted into nonperiodic form by taking second-order Taylor series expansions, as listed in the Appendix. The integral is then obtained analytically. Limits are taken such that sources always lie outside the enclosed region.

The system of equations for all b_i as described by (16) can be arranged into a matrix format,

$$A\sigma = b \qquad (17)$$

from which the values for σ's—the unknowns—are obtained with a suitable technique such as Gaussian elimination. The size of A is N by N for the oxidant diffusion and $2N$ by $2N$ for the oxide flow. Equation (16) is used again, this time with known σ's and different operators, to obtain boundary parameters desired. Note that we are free to choose any operators as long as they are reasonable. Our technique thus lends itself to extracting boundary parameters such as oxidant flux and shear stress at the oxide-silicon interface.

The freedom to define operators for extracting parameters from the intermediary solution is a unique feature of our approach. In contrast, the choice of parameters that can be obtained from Green's third identity (Green's formula), which is used in [1] and [2], is very limited. Granted Green's third identity does not require intermediary solutions, that is the sources, and thus is probably more accurate; ours, on the other hand, is more flexible.

A first-order method is used in the time-stepping scheme. After growth rate is computed from the solution for the oxidant diffusion, the oxide flow problem is solved to determine the velocities of the boundary segments. The boundary segments are then moved according to the time-step size chosen. This cycle of operations is repeated.

Applications

We now discuss a few examples on how to use our method. These technological applications include simulation of LOCOS processes, analysis of stress associated with deformation, and optimization of ambient pressure to reduce oxidation time.

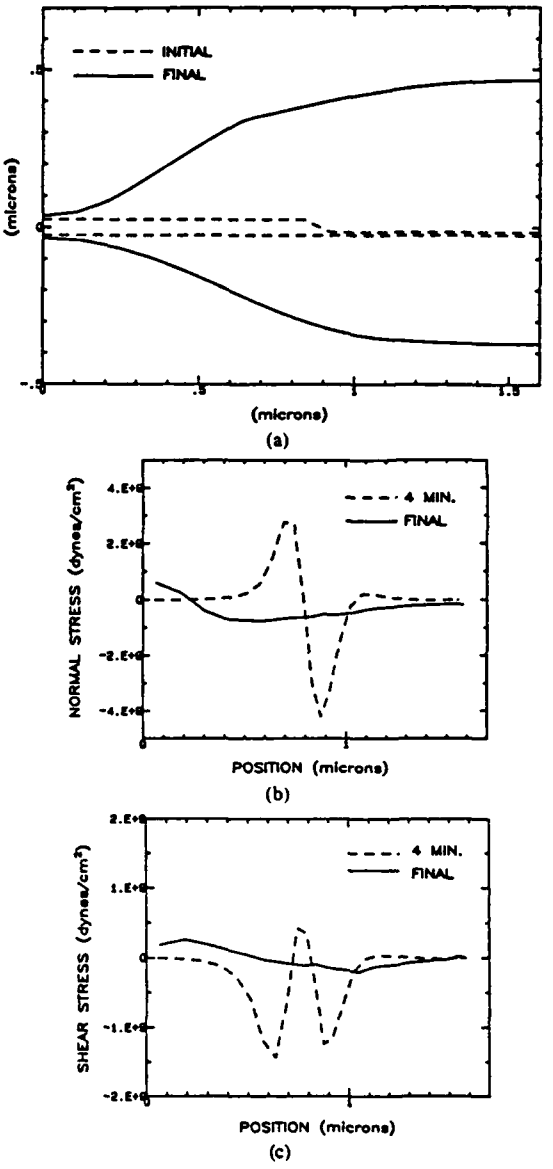


Fig. 2. Simulation of a SemiRox structure with pad oxide (relief oxide) of 500 Å. Final field oxide thickness is approximately 8000 Å. (a) Outline of the oxide layer before and after oxidation. (b) Normal stress distribution at the oxide-silicon interface. Note that the positive direction is into the oxide bulk. (c) Shear stress distribution at the oxide-silicon interface. Direction of force is tangent to the interface.

A. LOCOS Simulations

The two LOCOS presented are the semi-recess oxidation (SemiRox) and sidewall masked isolation (SWAMI) techniques.

The SemiRox structure has a pad oxide of 500 Å and a thin silicon nitride mask. As discussed earlier, the mask is assumed to be sufficiently thin so that is does not generate significant forces when bent. Oxidation is carried out at 1025°C in wet O_2. The shape of the oxide layer is shown in Fig. 2(a) with normal and tangential components of the stress distribution along the oxide-silicon interface given in Fig. 2(b) and (c), respectively. The value for the

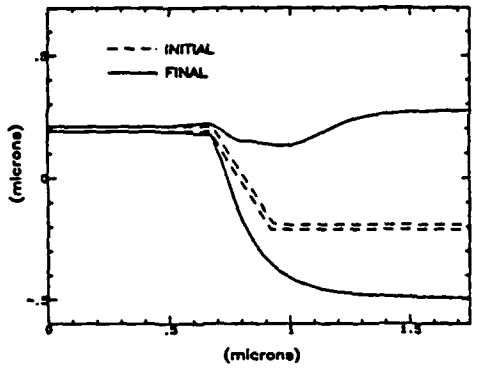


Fig. 3. Sloped-wall SWAMI oxidation. Pad oxide is 200 Å.

oxide viscosity is taken from [9]:

$$\mu(T) = \mu_0 \exp\left(\frac{E_\mu}{kT}\right) \tag{17}$$

where $\mu_0 = 1.586 \times 10^{-10}$ poises, $E_\mu = 5.761$ eV, k is the Boltzmann constant, and T is the temperature. The normal stress distribution is dipole due to the highly nonuniform oxide growth rate at the edge of the silicon nitride mask. Whether the stress is compressive (negative) or tensile (positive) depends on the oxide is being slow down or pulled. As the oxide layer gets thicker, the growth rate decreases and becomes more uniform, hence the decrease in the normal stress. Being more complicated, the behavior of the shear stress cannot be described in intuitive terms.

For the SWAMI process, we simulate an improved technique that uses sloped wall [10]. Simulation parameters are 200 Å pad oxide, 975°C, and wet O_2. The sidewall is sloped at 54.74°; silicon nitride covers the top and sidewall, and extends slightly into the trench. Fig. 3 shows the initial and final shapes of the structure.

In general, our results from different simulations agree well Chin's. The discrepancy (approximately 20 to 30 percent) in the normal stress is mainly due to the different ways of calculating stress—Chin approximated stress with hydrodynamic pressure whereas we dealt with the full stress tensor.

B. Deformation Stress Calculation

The plastic yield strength of single-crystal silicon is a strong function of temperature. At room temperatures, single-crystal silicon is known to fracture but never deform when subjected to stress [11]. Enhanced oxidation notwithstanding, defects do not appear in plasma-assisted oxidation, which is carried out at moderately elevated temperatures (between 300°C and 600°C) [12]. It is mainly in conventional high-temperature thermal oxidation that defects are observed.

Slip is the major defect species, its occurrence is confined to the {111} planes. On a {100} wafer typically used in making MOS devices, the {111} planes intersect on the surface at right angle relative to one another. Devices are usually aligned with these "fault lines" for ease of dicing.

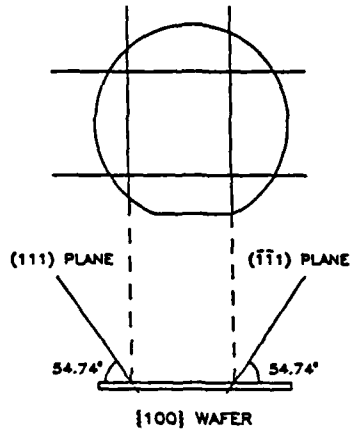


Fig. 4. Two slip planes on a {100} wafer, from [13].

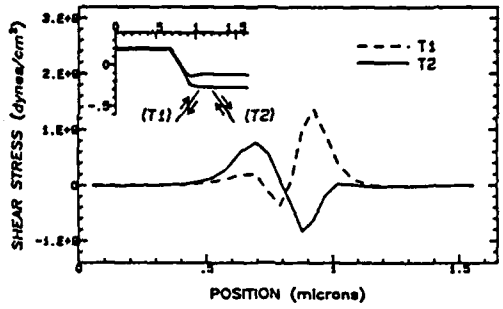


Fig. 5. Shear stress in slip planes. The shape of the oxide layer is shown in the insert, together with the slip directions. Simulation structure is the sloped-wall SWAMI oxidation given in Fig. 3.

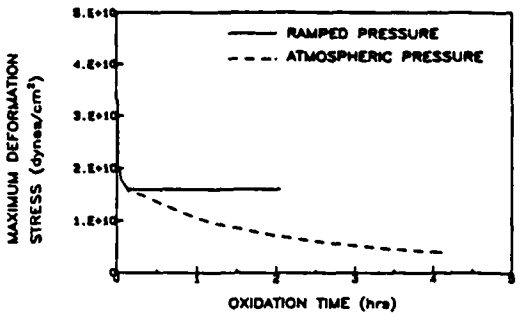
If we look at a suitable cross section of a device, two {111} planes make 54.74° with the surface, as shown in Fig. 4. Slip occurs when the shear stress on these slip planes exceeds the plastic yield strength of silicon.

The advantage of source distribution technique is apparent now. With the solutions defined by the intermediary sources, the normal stress and the shear stress tangent to the oxide–silicon interface can be determined. The same sources can also be used to calculate the shear stress in the slip planes. This is done by redefining the local coordinate $\eta - \xi$ of the shear stress operator to align with the slip planes. Fig. 5 shows the deformation stress associated with the two slip planes for the SWAMI structure mentioned easier.

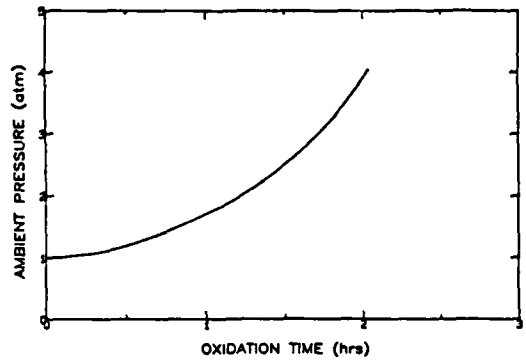
Strictly speaking, we cannot accurately model deformation shear stress in two dimensions; the reason has to do with the slip directions. Consider one of the {111} plane shown in Fig. 4. There are three ⟨110⟩ slip directions with this slip plane, one normal to the paper, the other two intersecting the paper at a low angle (30°). Our method can only approximate the shear stress associated with the latter two, while ignoring that of the first one.

C. Ramped-Pressure Oxidation

In the third application the feasibility of reducing oxidation time by ramping the ambient pressure is investigated. To produce the required oxide thickness, the field



(a)



(b)

Fig. 6. (a) Peak stress distribution (in slip planes) of conventional and ramped-pressure oxidation, plotted as a function of time. Simulation structure is similar to the SemiRox oxidation shown in Fig. 2. (b) Ambient pressure curve for ramped-pressure oxidation.

oxidation step must subject a wafer to high temperatures for long periods, thus causing significant impurity redistribution in the substrate. Knowing that stress is proportional to oxidation rate and that oxidation rate decreases as oxide layer gets thicker, we can attempt to shorten the oxidation time by gradually ramping the ambient pressure beyond atmospheric. In doing so we may however generate a stress that exceeds the plastic yield strength of silicon, hence damaging the substrate. The plastic deformation can be avoided if we increase the ambient pressure in such a way that the maximum stress always stays below the critical limit.

The test structure used in this study is SemiRox, with pad oxide of 500 Å. Oxidation is carried out at 950°C in wet O_2; the final field oxide thickness is approximately 8000 Å. A simulation is first done at constant atmospheric pressure to decide, in an arbitrary manner, a safe value for the maximum shear stress that the silicon substrate may withstand in its slip planes. This value is then used as a ceiling in a second simulation which ramps the ambient pressure. Our result indicates that the oxidation time can be reduced by half, roughly. Fig. 6(a) shows the maximum shear stress as a function of time for the case when ambient pressure is maintained at atmospheric and for the case when it is ramped. A plot of the ambient pressure in the ramped-pressure oxidation is given in Fig. 6(b), the pressure begins at 1 atm and increases to 4 towards the end.

Conclusion

A boundary integral equation method has been developed to model oxidation diffusion in two dimensions. It is a simple technique due to the fact that the oxidation diffusion can be modeled as a harmonic problem and the oxide motion as a biharmonic problem. In theory, problems such as the effect of stress on surface reaction rate can be readily incorporated into our method as long as they can be treated as a change in the boundary conditions.

Appendix
Boundary Parameters and Operators for Creeping Flow

Boundary parameters that may be obtained from the stream function are as follows:

(1) Normal component of flow velocity:

$$u_\eta = -\frac{\partial \Psi}{\partial \xi}. \tag{A1}$$

(2) Tangential component of flow velocity:

$$u_\xi = \frac{\partial \Psi}{\partial \eta}. \tag{A2}$$

(3) Vorticity:

$$\omega = -\nabla^2 \Psi. \tag{A3}$$

(4) Hydrodynamic pressure:

$$P = -\mu \bar{\nabla}^2 \Psi. \tag{A4}$$

(5) Normal stress:

$$T_{\eta\eta} = -P + 2\mu \frac{\partial u_\eta}{\partial \eta} = \mu \left[\bar{\nabla}^2 - 2 \frac{\partial^2}{\partial \eta \partial \xi} \right] \Psi. \tag{A5}$$

(6) Shear stress:

$$T_{\eta\xi} = \mu \left(\frac{\partial u_\eta}{\partial \xi} + \frac{\partial u_\xi}{\partial \eta} \right) = \mu \left[\frac{\partial^2}{\partial^2 \eta} - \frac{\partial^2}{\partial^2 \xi} \right] \Psi \tag{A6}$$

where $\hat{\eta}$ is unit normal vector pointing into the oxide bulk and $\hat{\xi}$ is unit tangent vector 90° clockwise from $\hat{\eta}$.

Boundary Conditions for LOCOS

At the free surface the oxidant concentration is maintained at the equilibrium concentration C^*; both the normal and shear stress components are zero since there is no force exchange between the oxide and the gas:

$$C = C^*$$
$$T_{\eta\eta} = 0$$
$$T_{\eta\xi} = 0. \tag{A7}$$

At the oxide–silicon interface the oxidation rate is proportional to both the oxidant concentration and normal component of oxidant flux, thus the following form of boundary condition is applied:

$$D\frac{\partial C}{\partial \eta} = k_s C \tag{A8A}$$

where k_s is the chemical surface reaction rate. The oxide–silicon interface recesses into the substrate at a velocity

$$\vec{u} = -\frac{k_s C}{\alpha N_1}\hat{\eta} \tag{A8B}$$

while oxide is injected into the bulk at

$$\vec{u} = (\alpha - 1)\frac{k_s C}{\alpha N_1}\hat{\eta}. \tag{A8C}$$

α is the volume of oxide produced from a unit volume of silicon, it is approximately 2.24. N_1 is the number of oxidant molecules required to form a unit volume of oxide.

At the oxide–silicon–nitride interface, the oxidant flux has a zero normal component:

$$\frac{\partial C}{\partial \eta} = 0. \tag{A9A}$$

The silicon nitride mask is assumed to be sufficiently thin such that is does not exert normal stress on the oxide:

$$T_{\eta\eta} = 0 \tag{A9B}$$

while maintaining a no-slip boundary condition:

$$u_\xi = 0. \tag{A9C}$$

Kernels

Among the 3 periodic kernels, only $\tilde{K}_2$ cannot be expressed in simple functions, but its derivatives can.

$$\tilde{K}(\vec{r}) = \tfrac{1}{2}\log\,[2(\cosh(y) - \cos(x))]$$

$$\tilde{K}_1(\vec{r}) = \tfrac{1}{2}y\log\,[2(\cosh(y) - \cos(x))] = y\tilde{K}(\vec{r})$$

$$\frac{\partial}{\partial x}\tilde{K}_2(\vec{r}) = \tilde{K}(\vec{r}) - y\frac{\partial}{\partial y}\tilde{K}(\vec{r})$$

$$\frac{\partial}{\partial y}\tilde{K}_2(\vec{r}) = y\frac{\partial}{\partial x}\tilde{K}(\vec{r}). \tag{A10}$$

The corresponding nonperiodic Taylor series approximations are obtained by expanding the trigonometric and hyperbolic functions to second order at $\vec{r} = 0$.

$$K(\vec{r}) = \tfrac{1}{2}\log\,(x^2 + y^2)$$

$$K_1(\vec{r}) = \tfrac{1}{2}y\log\,(x^2 + y^2)$$

$$K_2(\vec{r}) = \tfrac{1}{2}x\log\,(x^2 + y^2) - x. \tag{A11}$$

Acknowledgment

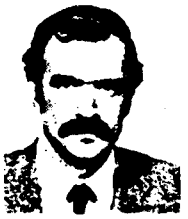
The authors would like to thank Dr. Ali Argon for discussions on plastic deformation of silicon and Dr. Daniel Meiron on integral equation methods.

References

[1] D. Chin, S. Y. Oh, and R. Dutton, "A general solution method for two-dimensional non-planar oxidation," *IEEE Trans. Electron Devices*, vol. ED-30, p. 993, 1983.
[2] H. Matsumoto and M. Fukuma, "A two-dimensional Si oxidation model including elasticity," *IEDM Tech. Dig.*, vol. 39, 1983.
[3] B. E. Deal and A. S. Grove, "General relationship for the thermal oxidation of silicon," *J. Appl. Phys.*, vol. 36, p. 3770, 1965.
[4] E. P. EerNisse, "Viscous flow of thermal SiO_2," *Appl. Phys. Lett.*, vol. 30, p. 290, 1977.
[5] W. E. Langlois, *Slow Viscous Flow*. New York: Macmillan, 1964.
[6] M. A. Jaswon and G. T. Symm, *Integral Equation Methods in Potential Theory and Elastostatics*. London, England: Academic, 1977.
[7] G. R. Baker, D. I. Meiron, and S. A. Orszag, "Generalized vortex methods for free-surface flow problems," *J. Fluid Mech.*, vol. 123, p. 477, 1982.
[8] E. Isaacson and H. B. Keller, *Analysis of Numerical Methods*. New York: Wiley, p. 316, 1966.
[9] D. Chin, S. Y. Oh, S. M. Hu, R. W. Dutton, and J. L. Moll, "Two-dimensional oxidation," *IEEE Trans. Electron Devices*, vol. ED-30, p. 744, 1983.
[10] K. Y. Chiu, J. L. Moll, K. M. Cham, J. Lin, C. Lage, S. Angelos, and R. L. Tillman, "The sloped-wall SWAMI—A defect-free zero bird's beak local oxidation process for scale VLSI technology," *IEEE Trans. Electron Devices*, vol. ED-30, p. 1506, 1983.
[11] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1976, p. 565.
[12] A. Riesman, private communication.
[13] S. K. Ghandi, *VLSI Fabrication Principles*. New York: Wiley, 1983, p. 11.

Thye-Lai Tung (S'85) was born in Malaysia. He received the B.S. and M.S. degrees in electrical engineering and computer science from Massachusetts Institute of Technology, Cambridge, in 1981 and 1984, respectively.

His research work is in the area of integrated circuit process simulation. He is currently working toward the Ph.D. degree.



Dimitri A. Antoniadis was born in Athens, Greece, in 1947. He received the B.S. degree in physics from the National University of Athens in 1970, and the Ph.D. degree in electrical engineering in 1976 from Stanford University, CA.

From 1970 to 1978 he was a Research Associate and Instructor in the Department of Electrical Engineering at Stanford University. He joined the faculty at MIT, Cambridge, in 1978 where he is at present an Associate Professor of Electrical Engineering. He is also Director of the new MIT Microsystems Laboratory. From 1969 to 1976, he conducted research in the area of measurements, modeling, and numerical simulation of the earth's ionosphere and thermosphere. He has designed and developed various electronic instruments, currently used for the study of the ionosphere at several locations throughout the world. His present scientific activity lies in the area of semiconductor device and integrated circuit technology. Until 1978, he directed the development of SUPREM (Stanford University Process Engineering Models). His research is currently focused on three areas: (a) fabrication and modeling of ultra-small geometry semiconductor devices; (b) silicon process modeling, with emphasis on the physics of processes and on on computer aids for process synthesis; and (c) silicon films on amorphous substrates with ultimate goal 3-dimensional integrated devices. He is author or coauthor of over 50 technical articles in the above areas.

Dr. Antoniadis is a member of the Electrochemical Society and of the Materials Research Society.

January, 1986 LIDS-P-1533

Dynamic Production Scheduling in Computer-Aided Fabrication of Integrated Circuits

by

Christopher Lozinski

and

Stanley B. Gershwin
617-253-2149

Massachusetts Institute of Technology

Laboratory for Information and Decision Systems

Cambridge, Massachusetts 02139

This paper will be presented at the IEEE Robotics and Automation Conference, April 7-10, 1986, San Francisco, California.

ABSTRACT

The scheduling of a VLSI manufacturing facility is being studied using MIT's new Integrated Circuits Laboratory (ICL) as a testbed. This paper describes the facility, outlines the scheduling methodology that is under investigation and presents the design of the initial preliminary scheduling system.

INTRODUCTION

Scheduling is the process of allocating the use of a set of scarce resources over time so as to best meet a set of objectives. In this case, the resources are the machines, technicians, and support systems of a semiconductor fabrication facility. The objectives are to satisfy the needs of the users of the facility with minimal delay. Because the resources are finite, there are times when users' needs may conflict. There must be a set of policies to deal with these conflicts, and the scheduler must resolve them in a way that is consistent with system policy.

THE MIT INTEGRATED CIRCUITS LABORATORY

The MIT Integrated Circuits Laboratory (ICL) is a new facility for semiconductor research. It will also serve as both a customer and laboratory for scheduling algorithm development. The building contains offices, clean rooms, laboratory space, and maintenance and support areas. Clean rooms have low particle counts due to constant flow of filtered air. Clean rooms contain either communal or private equipment. Laboratories are not necessarily located in clean rooms.

House systems are threaded through the building providing services such as air flow, humidity control, gaseous nitrogen, deionized water, and solvent waste removal.

A network of computers and terminals provides the hardware needed by the Computer-Aided Fabrication (CAF) project. A goal of this project is the efficient use of computers in semiconductor fabrication. As a part of the CAF project, scheduling is under study for both coordinating the activities in the ICL (to make the best possible use of valuable resources such as machines, technicians, power, steam, chilled water, and chemicals) and to develop algorithms for industry. The purpose of this paper is to describe our preliminary work in scheduling the ICL.

The ICL facility is run by a team of professionals. Technicians maintain and operate equipment. Process engineers are responsible for groups of machines and the processes that run on those machines. A manager is responsible for the smooth operation of the facility.

The facility is used by about 100 students and 20 faculty members whose interests include submicron structures, semiconductor materials, semiconductor devices, semiconductor processes, and CAF research.

Most research involves performing operations on silicon wafers. A group of wafers that is processed together is called a lot. The sequence of activities required to produce a lot is called a flow. A linear sequence of operations is called a linear flow. A flow can branch. Lots can be split for rework or experimentation. The flow of a research lot may also depend on measurements made after a process step. Wafers are discarded if unusable.

In addition to production and laboratory operations, other activities such as maintenance, training, inspection, repair, and machine setups must be scheduled.

PROBLEM DESCRIPTION

The scheduling problem is to choose a vector T where Tj is the time that activity number j is scheduled to begin. Tj is selected from the set of legal times for activity j. This set is limited due to personal schedule constraints, machine repair states, and previously scheduled activities such as preventive maintenance. T is further constrained by limited capacities of machines, house systems, and technicians, and by radio interference among machines and by conflicting setup requirements.

METHODOLOGY

The methodology selected for studying the scheduling of this facility involves the following:

1) Study semiconductor fabrication technology and the ICL to determine important features of the problem. The schedule is constrained by the structure of the system. For example:

a) Exposing wafers to air can dramatically reduce yield. Therefore certain sequences of activities must not have any delay.

b) The house system which supplies deionized water has a 300 gallon storage capacity, and a 30 gallon per minute generation capacity. A single wet sink uses 20 gallons per minute. An effective schedule must limit use of deionized water to within capacity.

c) The ion implanter has two different types of setups. The source and the wafer holder can both be changed. The scheduling system must insure that both setups are appropriate, or it must reserve machine and technician time to change them.

2) Specify objectives of scheduling

Possible scheduling goals include minimizing cycle time (the time a wafer spends in the system), maximizing throughput, minimizing cost, or satisfying a preferred customer. Low cycle time requires small buffers, high throughput requires larger buffers.

Lower costs means things should wait until several batches can be processed together. Satisfying a preferred customer means certain lots move first no matter what the consequences. Each of these objectives can lead to a different schedule.

The initial project goal is to provide a simple, effective scheduling system for the MIT ICL Later versions of the scheduling algorithm will meet criteria such as those described above.

To use optimal control techniques to help develop feedback rules for scheduling requires a numerical objective function L. One possible objective is the minimization of the average expected time between the arrival of an order for a part and the completion of the part.

3) Develop a solution technique.

The scheduling approach will follow the framework described in Gershwin et al. (1986). In that framework, the capacity constraints of the manufacturing system are explicitly modeled and feedback control strategies are developed to respond to random, potentially disruptive events

Because of the complexity of the problem, hierarchical scheduling methods are being developed. The top level of the hierarchy deals with relatively long term issues such as maintenance frequencies, setup frequencies, historical demand, etc. It allocates time on each of the machines for each class of activities. Since there is a substantial setup time between certain activities (such as when there are two different kinds of gases or impurities in a chamber), the frequencies of the changeovers must be chosen with care.

For example, given the best current estimates about usage during a given future time period (such as a week three weeks hence) the top level of the scheduler may decide that the ion implanter will be set up with the hearth stage from 9:00 AM Monday until 11:00 AM Wednesday, after which it will be changed over to the cold stage. Maintenance may then be scheduled until 2:00 PM, followed by a setup with the hearth stage. These allocations are chosen on the basis of historical data, without reference to current requests for that machine during that time period. In addition, some amount of time is built into the setup intervals to allow for such disruptions as machine failures.

As actual demand data arrives, the lowest level of the hierarchical scheduler assigns requests to these time intervals. For example, the period from 9:00 Monday until 11:00 Wednesday of that week is only used to satisfy requests for the ion implanter with the hearth stage. If the demand for the ion implanter with the hearth stage is greater than the allocated time, or if some users have schedule conflicts that prevents their use during that period, the excess is moved to the interval starting at 2:00 Wednesday.

The middle level of the scheduler has two functions: first, it responds to random events such as machine failures. When a machine goes down, it adjusts the schedule in accordance with facility policies about fairly allocating the inconvenience caused by the failure among the users that are affected. Second, it adjusts the setup intervals. If, for example, during some week there is greater than usual demand with the hearth stage and less for the cold stage, the 11:00 Wednesday changeover may be deferred until noon.

If there is a systematic discrepancy between the prescheduled setup intervals and the demand, it is the responsibility of the highest level of the algorithm to adjust the lengths of the intervals.

This preliminary statement is far from complete. We anticipate using methods such as those of Kimemia and Gershwin (1983) and Gershwin, Akella, and Choong (1985) to respond to machine failures, and Graves et al. (1983) and Watts (1986) for the detailed dispatching of lots.

4) Develop an appropriate database.

The detailed schedule is strongly dependent on a wide variety of kinds of information. Required data structures must represent flows, personal schedules, machine schedules, lot schedules, reservations, machine maintenance, machine setup states, machine repair states, the effects of failures of specific house systems, the maintenance of house systems, the priorities of lots, the facilities authorization list, and the skills list.

5) Test the entire system

The scheduling algorithms developed will be used to run the ICL as an experimental test of these ideas.

VERSION 1 SCHEDULING SYSTEM

As a first step towards solving the problem, a simple scheduling system is being implemented. This system is intended to be little more than a computerized sign-up sheet. Its objectives are to

a) Allow users to reserve equipment.

b) Inform users when machines are down, scheduled for maintenance, or reserved by someone else.

The table containing the desired information is shown in Fig 1. There is one column for every resource. The rows correspond to time slots. Future schedules are listed below the current line, previous ones above. The user sees a window into this table as shown in Fig. 2. Because of limited screen size, only four columns, representing four resources, are shown at a time.

CONCLUSION

Scheduling VLSI facilities is a new and important field. Competition is pushing American manufacturers away from mass production and into Application-Specific Integrated Circuits (A.S.I.C.). The rising cost of new facilities is putting pressure to process research lots on the same line as production lots. Both of these trends increase the number of flows in a single facility. More flows make scheduling more difficult and thus increase the opportunity for profiting from scheduling. This will make VLSI scheduling an important activity in the near future.

REFERENCES

Gershwin, S. B., R. Akella, and Y. F. Choong (1985), "Short-Term Production of an Automated Manufacturing Facility," IBM Journal of Research and Development, Vol. 29, No. 4, pp 392-400, July, 1985.

Gershwin, S. B., R. Hildebrant, R. Suri, and S. K. Mitter (1986), "A Control Perspective on Recent Trends in Manufacturing Systems," IEEE Control Systems Magazine, to appear.

Graves, S. C., H. C. Meal, D. Stefek, and A. H. Zeghmi (1983), "Scheduling of Re-Entrant Flow Shops," Journal of Operations Management, Vol. 3, No. 4, pp 197-207, August, 1983.

Kimemia, J. G. and S. B. Gershwin (1983), "An Algorithm for the Computer Control of Production in a Flexible Manufacturing System," IIE Transactions, Vol. 15, No. 4, pp 353-362, December, 1983.

Watts, H. A. (1986) "Dispatching - The Critical Automation Link," presented at the National Bureau of Standards Symposium on Real-Time Optimization, January 21-22, 1986.

ACKNOWLEDGMENTS

Sponsored by the Defense Advanced Projects Agency and monitored by the Office of Naval Research under Contract N00014-85K-0213. Lozinski is supported by a grant from the Microelectronics and Computer Technology Corporation (MCC).

END

FILMED

5-86

DTIC